

## Descubrimiento de Conocimiento en Bases de Datos Utilizando Técnicas de Morfología Matemática Borrosa

Noé N. Frago y Ramón Fuentes-González

Universidad Pública de Navarra, Departamento de Automática y Computación,  
31006 Pamplona-España (e-mail: noe.frago@unavarra.es, rfuentes@unavarra.es)

---

### Resumen

En este artículo se analiza el efecto, la utilidad y la interpretación de los filtros asociados a la Morfología Matemática Borrosa en procesos de Descubrimiento de Conocimiento en Bases de Datos. En particular se estudian los operadores morfológicos erosión, dilatación, apertura, cierre, Top-Hat y Hit-or-Miss. Usando información de las bases de datos y relaciones binarias ordinarias como elementos estructurales, se implementan algunos de estos filtros. Con esta implementación se justifica que, con estos filtros morfológicos, pueden analizarse datos estructurados como tabla de registros, obteniéndose información útil no evidente. Finalmente se analizan diversas características de los operadores definidos, ilustrando los resultados con un ejemplo.

*Palabras clave: descubrimiento de conocimiento, matemática borrosa, filtros morfológicos, bases de datos*

## Knowledge Discovery in Databases Using Fuzzy Mathematical Morphology Techniques

### ABSTRACT

In this paper the effect and usefulness of some filters associated to Fuzzy Mathematical Morphology operators, erosion, dilation, opening, closing, Top-Hat and Hit-or-Miss are analyzed. We propose the interpretation of those operators in processes of Database Knowledge Discovery. Using database information and binary crisp relations as structuring elements, some of these filters are implemented. Also, and using the morphological filters, it is verified that it is possible to analyze information constructed in table of records, interpreted as fuzzy subsets, obtaining not evident useful information. Finally, some properties of the defined operators are analyzed, illustrating the results with an example.

*Keywords: knowledge discovery, fuzzy mathematical morphology, morphological filters, databases*

## INTRODUCCIÓN

En una primera etapa, una parte importante del esfuerzo de las ciencias de la computación se dedicó a la representación y almacenamiento de la información adquirida del mundo real, para posteriormente obtener conocimiento útil no evidente. Fruto de ese esfuerzo, existen en la actualidad numerosos modelos para el almacenamiento de la información, así como para su representación, entre los que cabe destacar el modelo relacional (Codd, 1970). Un problema común a la hora de representar información en una base de datos, es que dicha información pueda estar afectada de imprecisión e incertidumbre en su definición. Para afrontar esta situación surgió (Zadeh, 1965) la teoría de conjuntos borrosos. Esa misma problemática obligó a desarrollar nuevos modelos de bases de datos, como el Modelo FIRST para manejar bases de datos borrosas. En los últimos años, con enormes cantidades de información almacenada, se están desarrollando técnicas que permitan analizarla y obtener conocimiento útil, hasta ese momento desconocido. Esas técnicas engloban metodologías muy diversas, que se agrupan bajo el nombre de Minería de Datos y extracción de Conocimientos en Bases de Datos (Frawley et al., 1991). En el campo del procesamiento de imágenes, el término Morfología Matemática (Matheron, 1967; Serra, 1982) hace referencia a una teoría y a una metodología específica no lineal, basada fundamentalmente en operaciones de la teoría de conjuntos. Se han desarrollado estas técnicas para el procesamiento y análisis de imágenes digitalizadas, tanto para las imágenes binarias, como para las que poseen diferentes tonos de gris o para las imágenes en color, consideradas todas ellas como subconjuntos del plano digital  $Z^2$ .

Desde un punto de vista geométrico, la morfología matemática consiste en el análisis de las imágenes mediante otra imagen prefijada en forma, denominada elemento estructurante y que determina, para cada píxel de la imagen a procesar, la vecindad que condiciona su modificación. Con ese análisis se pretende la obtención de información relevante subyacente en esas imágenes digitalizadas. Los operadores básicos de la Morfología Matemática son la dilatación y la erosión. Un apartado fundamental de esta teoría es el posterior desarrollo de filtros morfológicos, como funciones que permiten resaltar o eliminar determinados elemento de las imágenes. Entre esos filtros, cabe destacar la apertura y el cierre. En Serra (1988), la Morfología Matemática se amplía a contextos más generales, relacionados con un tipo distinguido de conjuntos ordenados: los retículos completos (Heimans y Vincent, 1993). Posteriormente, se extiende esta teoría a subconjuntos borrosos (Bloch y Maître, 1995; Divyendu y Dougherty, 1992), aunque siempre desde la perspectiva del tratamiento de imágenes.

En un contexto distinto del tratamiento de imágenes digitalizadas, se introduce (Fuentes-González et al., 2002) la aplicación de técnicas de la Morfología Matemática borrosa en el campo de análisis de datos, posibilitando la selección de registros de una tabla  $U$  a partir de un subconjunto borroso  $A$  de  $U$  y de una relación  $R$  (borrosa o nítida) en  $U$ . La novedad de estas técnicas reside en que los criterios de selección se realizan fundamentalmente, teniendo en cuenta características de vecindad entre elementos, vecindad que se determina mediante la mencionada relación binaria  $R$ , que hace el papel de elemento estructurante. Los efectos de filtrado son el resultado de la composición de los operadores morfológicos básicos erosión y dilatación, que se construyen a utilizando t-normas e implicaciones borrosas. En este trabajo, por una parte, se extienden resultados anteriores (Fuentes-González et al., 2002 y Frago y Fuentes-González, 2006) y, por otra, se justifica posteriormente la teoría mediante su aplicación a un ejemplo real que proporciona una base de datos estructurada con 15.000 registros (apartado EJEMPLO de este trabajo).

## ANTECEDENTES: OPERADORES EROSIÓN Y DILATACIÓN BORROSA

La erosión y la dilatación son las operaciones básicas en la teoría Morfología Matemática, a partir de las cuales se definen la mayoría de los posteriores filtros utilizados. Comenzaremos aquí presentando la versión de estas definiciones en el contexto que nos interesa: el de descubrimiento de conocimientos en bases de datos (Fuentes-González, 2002; Frago 2006).

Sea  $U$  un referencial, conjunto finito cuyos elementos son registros en una base de datos.

Sea  $([0,1], \leq)$  la cadena completa  $[0,1] \subseteq R$  con la relación de orden usual  $\leq$ . Consideramos en  $[0,1]$  las siguientes operaciones:

(a) Una t-norma  $T: [0,1] \times [0,1] \rightarrow [0,1]$  tal que, para todo  $M \subseteq [0, 1]$  verifique:

$$T(a, \sup M) = \sup(T(a, M)).$$

(b) La implicación borrosa  $I: [0,1] \times [0,1] \rightarrow [0,1]$  obtenida por residuación (Erich et al., 2000) de la

t-norma anterior  $T$ , es decir,  $T: I(a, b) = \sup \{w \mid T(a, w) \leq b\} \quad \forall (a, b)$

Con estos elementos, fijada una relación borrosa  $R$  en  $U$   $R: U \times U \rightarrow [0,1]$  (que llamamos relación estructurante), y un subconjunto borroso  $A$  de  $U$ , Fuentes-González, R., et al. 2002, proponen la siguiente definición de operadores erosión y dilatación.

**Definición 1.-** Sea  $A: U \rightarrow [0,1]$  un subconjunto borroso de registros y sea  $R$  una relación estructurante. Se define la erosión  $E_R(A)$  y la dilatación  $D_R(A)$  de  $A$  asociadas a la relación estructurante  $R$  como los nuevos subconjuntos borrosos tales que,

$$E_R(A)(x) = \inf \{I(R^{op}(x,y), A(y)) \mid y \in U\} \quad \forall x \in U \quad y \quad D_R(A)(x) = \sup \{T(R(x,y), A(y)) \mid y \in U\} \quad \forall x \in U, \quad (1)$$

siendo  $T$  una t-norma del tipo (a),  $I$  su implicación residuada dada por (b) y  $R^{op}$  la relación borrosa opuesta de  $R$  definida por  $R^{op}(x,y) = R(y,x) \quad \forall (x,y) \in U \times U$ .

Si para  $x \in U$  se considera su  $R$ -vecindad, es decir el subconjunto borroso  $R_x$  tal que  $R_x(y) = R(x, y) \quad \forall y \in U$ , las expresiones (1) pueden representarse por:

$$E_R(A)(x) = \inf \{I(R_x^{op}(y), A(y)) \mid y \in U\} \quad \forall x \in U \quad y \quad D_R(A)(x) = \sup \{T(R_x(y), A(y)) \mid y \in U\} \quad \forall x \in U. \quad (2)$$

En las condiciones de la definición 1, si  $\leq$ ,  $\wedge$  y  $\vee$  representan, las extensiones puntuales habituales a  $[0,1]^U$  de la relación de orden y de las operaciones mínimo y máximo en  $[0,1]$ , se verifica el siguiente teorema.

**Teorema 1.-** Sean  $A$  y  $B$  subconjuntos borrosos de registros, y sea  $R$  una relación estructurante entre registros. Entonces los operadores erosión y dilatación morfológicas definidos en (1) verifican:

- 1)  $E_R(A \wedge B) = E_R(A) \wedge E_R(B) \quad \forall (A, B) \in [0,1]^U \times [0,1]^U$
- 2)  $D_R(A \vee B) = D_R(A) \vee D_R(B) \quad \forall (A, B) \in [0,1]^U \times [0,1]^U$
- 3) Si  $A \leq B$ , entonces,  $E_R(A) \leq E_R(B)$  y  $D_R(A) \leq D_R(B)$ .

Estos operadores nos permiten manipular los registros de una tabla  $U$  utilizando un subconjunto borroso  $A$  y una relación  $R$  (borrosa o no), obteniendo los nuevos subconjuntos borrosos  $E_R(A)$  y  $D_R(A)$ . El grado de pertenencia borrosa de un registro  $x$  de  $U$  a cada uno de estos nuevos subconjuntos borrosos se ha determinado, no por las características de dicho elemento, sino en función de las relaciones de sus  $R$ -vecindades  $R_x$  y  $R_x^{op}$  con los elementos del subconjunto borroso  $A$ .

Un caso particular interesante se obtiene, al considerar los operadores morfológicos asociados a la implicación de Lukasiewicz en  $[0,1]$ . Concretamente, considerando en  $[0,1]$  la t-norma determinada por la expresión:

$T(a, b) = \max(0, a+b-1) \quad \forall (a,b) \in [0,1] \times [0,1]$ , cuya implicación asociada es la de Lukasiewicz:  
 $I(a, b) = \min(1, 1-a+b) \quad \forall (a, b) \in [0,1] \times [0,1]$ . De esta forma las expresiones (1) para los operadores erosión y dilatación son ahora:

$$E_R(A)(x) = \min\{\min(1, 1-R(y,x)+A(y)) \mid y \in U\} \quad \forall x \in U$$

$$D_R(A)(x) = \max\{\max(0, R(x,y)+A(y)-1) \mid y \in U\} \quad \forall x \in U \quad (3)$$

El interés de este caso reside en que, al estar además relacionadas la t-norma y la implicación de Lukasiewicz por medio de la negación de Zadeh:  $1-T(a, 1-b) = I(a, b) \quad \forall (a, b) \in [0,1] \times [0,1]$ , se induce también una relación, entre los operadores erosión y dilatación morfológicos:

$$E_R(A') = D_R^{op}(A) \quad \forall A \in [0,1]^U, \quad (4)$$

en donde  $A'$  representa el subconjunto borroso  $A'(x) = 1-A(x) \quad \forall x \in U$ .

Aunque el elemento estructurante utilizado para calcular erosiones y dilataciones morfológicas puede ser cualquier relación borrosa  $R \in [0,1]^{U \times U}$ , hacemos notar que aparece otro caso distinguido cuando se utilizan como elementos estructurantes *relaciones nítidas*  $R$ ,  $R(U \times U) \subseteq \{0,1\}$ . En este caso  $R_x$  y  $R_x^{op}$  representan los subconjuntos ordinarios  $\{y \in U \mid R(x,y)=1\}$  y  $\{y \in U \mid R(y,x)=1\}$ . Además, en este caso, las expresiones (1) para la erosión y dilatación morfológicas no dependen de la t-norma y de la implicación, ya que las expresiones son:

$$E_R(A)(x) = \begin{cases} 1 & \text{si } R_x^{op} = \emptyset \text{ ó si } A(z)=1 \quad \forall z \in R_x^{op} \\ \inf\{A(z) \mid R(z,x)=1\} & \text{en otro caso} \end{cases} \quad \text{y} \quad D_R(A)(x) = \begin{cases} 0 & \text{si } R_x = \emptyset \text{ ó si } A(z)=0 \quad \forall z \in R_x \\ \sup\{A(z) \mid R(x,z)=1\} & \text{en otro caso} \end{cases} \quad (5).$$

Obsérvese que  $E_R(A)(x) = 0$  si y solo si existe  $z \in R_x^{op}$  tal que  $A(z) = 0$ , y que  $D_R(A)(x) = 0$  si y solo si  $R_x = \emptyset$  ó  $A(z) = 0$  para todo  $z \in R_x$ .

En estas condiciones, si para un  $x \in U$  se verifica que  $E_R(A)(x) = \alpha$  entonces la  $R^{op}$ -vecindad de  $x$  esta incluida en el  $\alpha$ -corte de  $A$ :  $R_x^{op} \subseteq A_\alpha$ . En el caso de la dilatación, si para un  $x \in U$  se verifica que  $D_R(A)(x) = \alpha$  entonces, para todo  $z$  perteneciente a la  $R$ -vecindad de  $x$  se verifica que  $A(z) \leq \alpha$ . En este caso además de los resultados que aparecen en el teorema 1, si la relación  $R$  es reflexiva y  $\subseteq$  representa la extensión puntual habitual de  $\leq$  a  $[0,1]^U$ , entonces  $E_R(A) \subseteq A \subseteq D_R(A)$ , lo que justifica los nombres de erosión y dilatación en función de sus efectos sobre  $A$ .

#### Filtros Morfológicos: apertura y cierre morfológicos

La composición de los operadores anteriores nos permite construir nuevos operadores, *filtros morfológicos*, sobre subconjuntos de registros. En particular, los filtros más simples son la apertura y el cierre. Con ellos, al igual que en el contexto del procesamiento de imágenes, se pueden construir nuevos filtros por unión e intersección de aquellos (en paralelo) o por composición (en serie) de los mismos. Definimos a continuación los operadores apertura y cierre.

**Definición 2.-** Sea  $A$  un subconjunto borroso de registros y  $R$  una relación borrosa entre registros. Se define la *apertura morfológica*  $\gamma_R$  y el *cierre morfológico*  $\zeta_R$  de  $A$  asociados a la relación estructurante  $R$  respectivamente por:

$$\gamma_R(A) = D_R(E_R(A)) \quad \text{y} \quad \zeta_R(A) = E_R^{op}(D_R^{op}(A)) \quad \forall A \in [0,1]^U. \quad (6)$$

La forma en que se construyen los operadores apertura y cierre determinan su función de filtro, la erosión ("con efecto eliminar") y la dilatación ("con efecto añadir") caracterizan los nuevos subconjuntos obtenidos  $\gamma_R(A)$  y  $\zeta_R(A)$ .

Si  $I$  es la implicación residuada asociada a la t-norma  $T$ , y  $E_R$  y  $D_R$  son los operadores morfológicos dados por (1), los nuevos operadores apertura  $\gamma_R$  y cierre  $\zeta_R$  verifican el siguiente teorema.

**Teorema 2.-** Sean  $A$  y  $B$  dos subconjuntos borrosos de la tabla de registros  $U$ , y sea  $R$  una relación binaria de  $U$ . Entonces los operadores apertura y cierre verifican:

- 1) Si  $A \leq B$ , entonces:  $\gamma_R(A) \leq \gamma_R(B)$  y  $\zeta_R(A) \leq \zeta_R(B)$ .
- 2)  $\gamma_R(A) \leq A \leq \zeta_R(A) \quad \forall A \in [0,1]^U$ .
- 3)  $\gamma_R(\gamma_R(A)) = \gamma_R(A)$ ,  $\zeta_R(\zeta_R(A)) = \zeta_R(A) \quad \forall A \in [0,1]^U$ .
- 4) Además, en el caso en que  $T$  sea la t-norma  $T(a, b) = \max(0, a+b-1) \quad \forall (a, b)$  o si la relación estructurante  $R$  es nítida, entonces:  $(\gamma_R(A))' = \zeta_R(A') \quad \forall A \in [0,1]^U$ .

En este caso, además de los resultados que aparecen en el teorema 2, si la relación  $R$  es reflexiva y  $\subseteq$  representa la extensión puntual habitual a  $[0,1]^U$ , entonces  $E_R(A) \subseteq \gamma_R(A) \subseteq A \subseteq \zeta_R(A) \subseteq D_R(A)$ .

Al igual que en el contexto del tratamiento de imágenes, podemos considerar nuevos operadores asociados a los anteriores, como los gradientes utilizados en el procesamiento de imágenes para la obtención de contornos. Ejemplo de esos gradientes son los de tipo Top-Hat.

**Definición 3.-** Sea  $A$  un subconjunto borroso de registros. Sea  $R$  una relación borrosa entre registros. Se define el gradiente *Top-Hat apertura*  $TH_{a,R}$  (respectivamente *Top-Hat cierre*  $TH_{c,R}$ ) de  $A$  asociado a  $R$  por:

$$TH_{a,R}(A)=A\wedge(\gamma_R(A))' \quad y \quad TH_{c,R}(A)=\zeta_R(A)\wedge A', \tag{7}$$

en donde  $\wedge$  representa la intersección de subconjuntos borrosos asociada a una t-norma. Se obtiene así los operadores *Top-Hat apertura*  $TH_a$  y *Top-Hat cierre*  $TH_c$  en  $[0, 1]^U$ .

EJEMPLO

A lo largo de este trabajo, utilizaremos el universo  $U$  formado por los registros de una tabla de acceso público (que se puede encontrar en la dirección [www.isc.uic.edu/~mlearn/MLSummary.html](http://www.isc.uic.edu/~mlearn/MLSummary.html)), que está relacionada con el censo de EE.UU. Dicha tabla tiene 15.000 registros y en ella se recogen datos relativos a personas, los campos de dicha tabla pueden verse en la Tabla 1.

En primer lugar, hemos construimos a partir de esa tabla un subconjunto borroso  $A$  de  $U$ , asociado a una función de grado de pertenencia  $A: U \rightarrow [0, 1]$ . Para ello, consideramos la función obtenida por el cociente  $SalAnuHts(x)=\frac{Ca3(x)}{Ca13(x)}$ , que llamamos salario anual por hora de trabajo semanal. Como  $Ca3(x)$  es el salario anual y  $Ca13(x)$  el número de horas de trabajo a la semana, el cociente puede interpretarse como una medida de la “bondad” del pago del trabajo. A continuación, utilizando este cociente y la constante  $SalAnuHtsMe$ , (salario anual por hora de trabajo semanal medio), obtenida mediante,  $SalAnuHtsMe = (\sum_{i=1}^{15000} SalAnuHts(x_i))/15000 = 5757$ . Definimos finalmente el subconjunto borroso  $A$ , mediante la función de pertenencia dada por la Formula 9. Cuyo grafo se representa en la Figura 1 (para el trazado de este grafo se ordenaron los registros por los valores de la variable  $SalAnuHts$ ). Evidentemente, el grafo de  $A(x)$ , depende de la definición de la función grado de pertenencia al subconjunto borroso  $A$ .

En la Tabla 2 muestran los valores de algunos estadísticos descriptivos de los conjuntos  $U$  y  $A$ , que se compararan posteriormente con los correspondientes a los de los subconjuntos borrosos  $E_R(A)$ ,  $D_R(A)$ ,  $\gamma_R(A)$ ,  $TH_{a,R}(A)$ ,  $TH_{c,R}(A)$  y  $HM_{(R1,R2)}(A)$ , que hemos tomado como referencias en este análisis. Son medias y porcentajes. Concretamente el cálculo de las medias se obtiene mediante la siguiente fórmula:

$$Media(P,Q)=\frac{\sum_{x\in Sop(Q)} P(x)}{Card(Sop(Q))}. \tag{8}$$

En donde  $P$  se sustituye por  $A$ ,  $Ca1$ ,  $Ca5$ ,  $Ca13$  y en donde  $Q$  se sustituye por los conjuntos,  $U$ ,  $A$ ,  $E_R(A)$ ,  $D_R(A)$ ,  $\gamma_R(A)$ ,  $TH_{a,R}(A)$ ,  $TH_{c,R}(A)$  y  $HM_{(R1,R2)}(A)$ . La expresión  $Sop(Q)$  representa el soporte del subconjunto borroso  $Q$  y  $Card(Sop(Q))$  indica el cardinal de  $Sop(q)$ . De forma similar se han calculado los porcentajes.

Para obtener conocimiento de  $U$  a partir de  $A$ , utilizaremos como relación estructurante la relación binaria ordinaria (nítida)  $R \subseteq U \times U$  definida por:

$$R(x, z) = 1 \Leftrightarrow (x.Ca1 = z.Ca1 \text{ AND } |x.Ca5-z.Ca5| \leq 1 \text{ AND } x.Ca9 = z.Ca9 \text{ AND } x.Ca10 = z.Ca10). \text{ En otro caso } R(x, y) = 0.$$

Esta relación es reflexiva y simétrica, es decir, es una relación de semejanza (o tolerancia). Y expresa lo siguiente: “dos personas  $x$  y  $z$  están relacionadas, si y sólo si  $x$  y  $z$  tienen la misma edad y sexo, son de la misma etnia y su diferencia de años de estudios es menor o igual a 1”. En este caso las personas  $x$  e  $y$  son semejantes.

Se puede considerar que  $A(x)$  nos da una medida de la proximidad, por exceso o por defecto, de  $SalAnuHts(x)$  respecto de la constante  $SalAnuHtsMe$ .

$$A(x)=\begin{cases} 0 & \text{si } SalAnuHts(x)<2600 \\ \frac{SalAnuHts(x)-2600}{5500-2600} & \text{si } 2600\leq SalAnuHts(x)\leq 5500 \\ 1 & \text{si } 5500<SalAnuHts(x)<6014 \\ \frac{8904-SalAnHts(x)}{8914-6014} & \text{si } 6014\leq SalAnuHts(x)\leq 8914 \\ 0 & \text{si } 8914<SalAnuHts(x) \end{cases} \tag{9}$$

Tabla 1: Campos de la Tabla U

Ca1	Edad	Ca8	Parentesco
Ca2	Tipo de empleo	Ca9	Etnia
Ca3	Sueldo	Ca10	Sexo
Ca4	Estudios	Ca11	Ganancia
Ca5	Años de estudio	Ca12	Pérdida
Ca6	Estado civil	Ca13	Horas de trabajo a la semana
Ca7	Empleo	Ca14	País de nacimiento

Tabla 2: Estadísticos descriptivos de U y A

	U	A
Media GraPerA	0.4	0.6
Media de edad	38.5	39
Media años-estudio	10.1	10
Media H-T-S	40.5	41.2
Media SalAnu	189980	199839
Porcentaje de mujeres	33%	32,8%
Porcentaje de blancos	85.8%	85,9%

ANÁLISIS DE LOS SUBCONJUNTOS:  $E_R(A)$ ,  $D_R(A)$ ,  $\gamma_R(A)$ ,  $\zeta_R(A)$ ,  $TH_{a,R}(A)$  Y  $TH_{c,R}(A)$

En adelante, las gráficas de las distintas figuras se corresponden con los datos del ejemplo anterior y con los registros de  $U$  ordenados en función de la variable  $SalAnuHts$ . Los resultados de los teoremas y proposiciones siguientes son independientes de los datos del ejemplo, pero se utilizan para ilustrar los resultados. Las siguientes propiedades nos aportan información sobre distintos agrupamientos de datos:

- Proposición 1.- Sea  $x \in U$  y  $R$  una relación reflexiva y simétrica (semejanza), entonces:
- 1)  $E_R(A)(x) = \alpha$  si y sólo si existe  $z_1$  semejante a  $x$  tal que  $A(z_1) = \alpha$  y para todo  $z$  semejante a  $x$  se verifica que  $A(z) \geq \alpha$ .
  - 2)  $D_R(A)(x) = \beta$  si y sólo si existe  $z_2$  semejante a  $x$ , tal que  $A(z_2) = \beta$  y para todo  $z$  semejante a  $x$  se verifica que  $\beta \geq A(z)$ .
  - 3) Si  $x \in U$  es tal que  $E_R(A)(x) \geq \alpha$  y  $D_R(A)(x) \leq \beta$ , entonces  $\alpha \leq A(z) \leq \beta \ \forall z \in R_x$ .
  - 4)  $\alpha = \beta = A(x)$  si y sólo si  $A(z) = A(x) \ \forall z \in R_x$ .
  - 5)  $E_R(A)(x) = A(x)$  si y sólo si  $R_x = \{x\}$  ó  $A(z) \geq A(x) \ \forall z \in R_x$
  - 6)  $E_R(A)(x) = 0$  si y sólo si existe  $z$  semejante a  $x$ , tal que  $A(z) = 0$ .
  - 7)  $E_R(A)(x) = 1$  si y sólo si  $A(z) = 1$  para todo  $z$  semejante a  $x$ .
  - 8)  $D_R(A)(x) = A(x)$  si y sólo si  $R_x = \{x\}$  ó  $A(z) \leq A(x) \ \forall x \in R_x$ .
  - 9)  $D_R(A)(x) = 0$  si y sólo si  $A(z) = 0$  para todo  $z$  semejante a  $x$ .
  - 10)  $D_R(A)(x) = 1$  si y sólo si existe  $z$  semejante a  $x$ , tal que  $A(z) = 1$ .

Según el tercer apartado de la proposición 1, los valores de la erosión y de la dilatación de un elemento  $x$  determinan el subintervalo  $[\alpha, \beta]$  de  $[0, 1]$ , en el que están los grados de pertenencia al subconjunto  $A$  de los semejantes a  $x$ . Esto permite determinar, el rango de los valores de  $SalAnuHts(y)$  con  $y$  semejante a  $x$ . Nótese que los intervalos determinados, con sus respectivos significados, se mantienen para los elementos del conjunto  $\{x \in U \mid E_R(A)(x) = \alpha, D_R(A)(x) = \beta\}$ . Es decir, dados  $\alpha$  y  $\beta \in [0, 1]$ , determinan un subconjunto  $H$  de  $U$ ,  $H = \{x \in U \mid E_R(A)(x) = \alpha, E_R(A)(x) = \beta\}$ , tal que, los semejantes de cualquier elemento de  $H$  tienen su grado de pertenencia al subconjunto  $A$  dentro del intervalo  $[\alpha, \beta]$ .

En la Tabla 3 se dan los datos descriptivos correspondientes a los conjuntos erosión y dilatación de  $A$  asociadas a la relación  $R$ , puede observarse diferencias apreciables respecto a los datos de  $A$  (Tabla 2) y entre los datos de  $E_R(A)$  y  $D_R(A)$ . En las Figuras 2 y 3 se representan los grafos de las funciones de pertenencia a los subconjuntos borrosos  $E_R(A)$  y  $D_R(A)$ . Queda reflejado que,  $E_R(A)(x) \leq A(x) \leq D_R(A)(x)$ , que  $E_R(A)(x) > 0$  solo para 658 personas ( $Sop(E_R(A))= 658$ ) y que  $D_R(A)(x) = 1$  para la mayoría registros (11826).

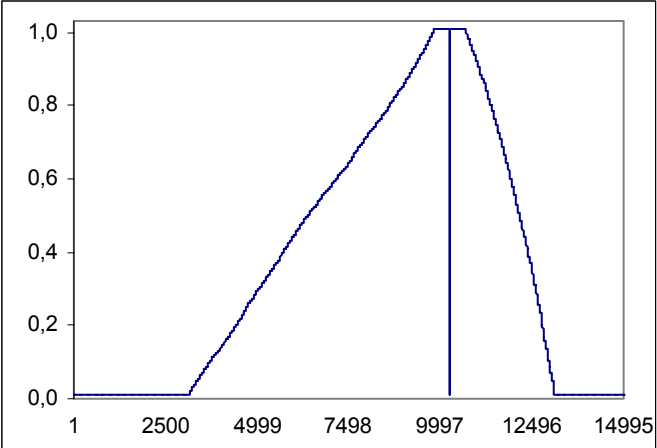


Fig. 1: Grafo de la función A(x)

Tabla 3: Estadísticos de $E_R(A)$ y $D_R(A)$		
	$E_R(A)$	$D_R(A)$
Media GraPerA	0.6	0.4
Media de edad	44,5	38,3
Media años de estudio	8.7	10.1
Media H-T-S	39.6	40.6
Media SalAnu	195595	190249
Porcentaje de mujeres	45.1%	32.7%
Porcentaje de blancos	29.6%	86.8%

En el caso de los operadores apertura y cierre asociados al subconjunto borroso A y a la relación R, se verifica las Proposiciones 2 y 3 que nos proporcionan nuevos elementos para el análisis de datos.

**Proposición 2.-** Sea  $x \in U$  y R una relación reflexiva y simétrica, entonces:

1)  $\gamma_R(A)(x) = \gamma$  si y sólo si  $\exists z \in R_x$  tal que  $A(h) \geq \gamma \ \forall h \in R_z$  y además  $\forall z \in R_x \ \exists h \in R_z$  tal que  $A(h) \leq \gamma$ .

2)  $\zeta_R(A)(x) = \mu$  si y sólo si  $\exists z \in R_x$  tal que  $A(h) \leq \mu \ \forall h \in R_z$  y además  $\forall z \in R_x \ \exists h \in R_z$  tal que  $A(h) \geq \mu$ .

En las Figura 4 y 5 se representan los grafos de las funciones de pertenencia a los subconjuntos borrosos  $\gamma_R(A)$  y  $\zeta_R(A)$ . Queda reflejado que  $\gamma_R(A)(x) \leq A(x) \leq \zeta_R(A)(x)$  y que en este caso, la mayoría de las personas toman valores cero para la apertura y uno para el cierre.

**Proposición 3.-** Sea  $x \in U$  y R una relación reflexiva y simétrica (semejanza). Entonces:

1)  $\gamma_R(A)(x) = A(x)$  si y sólo si  $\exists z \in R_x$  tal que  $A(h) \geq A(x) \ \forall h \in R_z$  y  $\forall z \in R_x \ \exists h \in R_z$  tal que  $A(h) \leq A(x)$ .

2)  $\zeta_R(A)(x) = A(x)$  si y sólo si  $\exists z \in R_x$  tal que  $A(h) \leq A(x) \ \forall h \in R_z$  y  $\forall z \in R_x \ \exists h \in R_z$  tal que  $A(h) \geq A(x)$ .

3) Si  $E_R(A)(x) = \alpha = \beta = D_R(A)(x) \Rightarrow \gamma = \mu = \alpha = \beta$ .

4)  $\gamma_R(A)(x) = 0$  si y sólo si  $\forall z \in R_x$  existe  $h \in R_z$  tal que  $A(h) = 0$

5)  $\zeta_R(A)(x) = 0$  si y sólo si existe  $z \in R_x$  tal que  $A(h) = 0 \ \forall h \in R_z$

6)  $\gamma_R(A)(x) = 1$  si y sólo si existe  $z \in R_x$  tal que  $A(h) = 1 \ \forall h \in R_z$

7)  $\zeta_R(A)(x) = 1$  si y sólo si  $\forall z \in R_x$  existe  $h \in R_z$  tal que  $A(h) = 1$

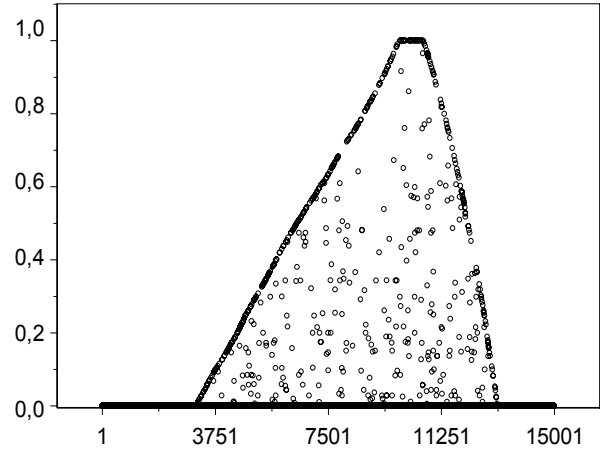


Fig. 2: Grafo de  $E_R(A)(x)$

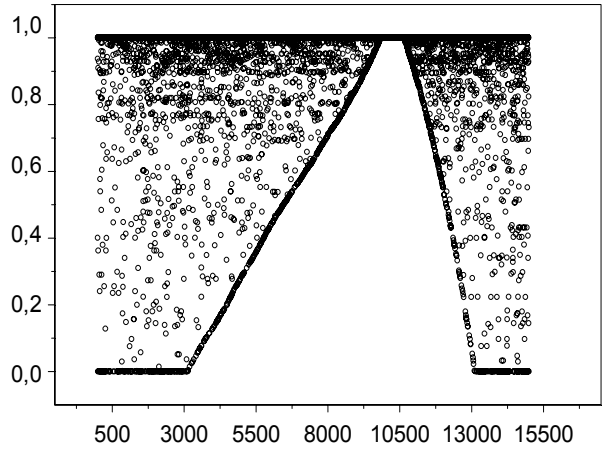


Fig. 3: Grafo de  $D_R(A)(x)$

En nuestro ejemplo, la Tabla 4 recoge datos descriptivos correspondientes al conjunto apertura de A asociada a la relación R. Pueden apreciarse diferencias apreciables respecto a los datos de A y una similitud respecto a los datos de  $E_R(A)$ .

Para el operador Top-Hat asociado al subconjunto borroso A y a la relación R, se verifica la siguiente proposición.

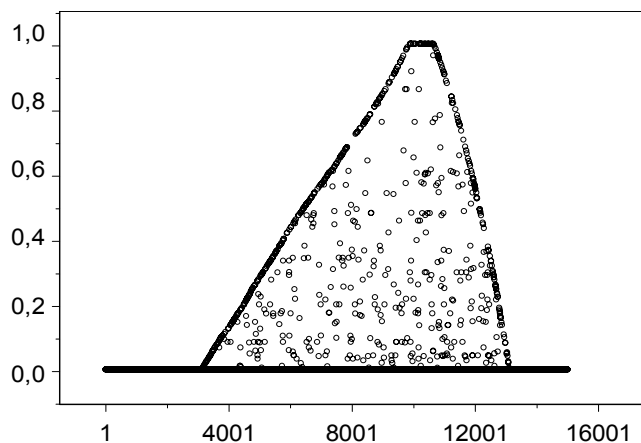


Fig. 4: Grafo de  $\gamma_R(A)(x)$

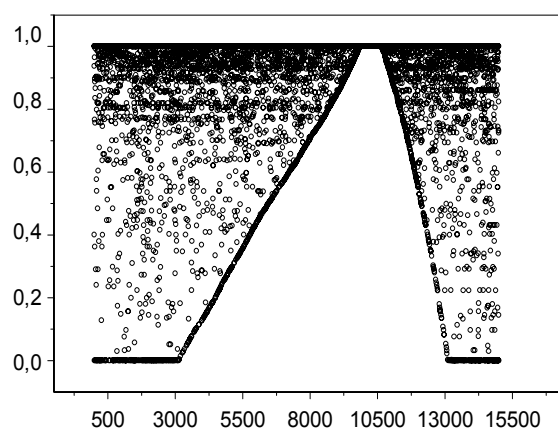


Fig. 5: Grafo de  $\zeta_R(A)(x)$

**Proposición 4.-** Sea  $x \in U$  y  $R$  una relación reflexiva y simétrica, entonces:

- 1) El Top-Hat de la apertura se puede escribir,  $TH_{a,R}(A)(x) = \min(A(x), 1 - \gamma_R(A)(x)) \quad \forall x \in U$ .
- 2)  $TH_{a,R}(A)(x) = \alpha$  para  $x \in U$  y  $\alpha \in [0, 1]$  es equivalente a:  $A(x) \geq \alpha$ , existe  $z \in R_x$  tal que  $A(h) \geq \gamma_R(A)(x)$   $\forall h \in R_z$  y  $\forall z \in R_x$  existe  $h \in R_z$  tal que  $A(h) \leq \gamma_R(A)(x)$ .

Obsérvese que  $\gamma_R(A)(x) \leq 1 - \alpha$ .

Decir que para  $x \in U$  se verifica que  $TH_{a,R}(A)(x) = 0.7$  es equivalente a decir que:  $A(x) \geq 0.7$ , que para todo  $z \in R_x$  existe  $h \in R_z$  tal que  $A(h) \leq 0.3$  y que existe un  $z \in R_x$  tal que  $A(h) \geq \gamma_R(A)(x) \quad \forall h \in R_z$  ( $\gamma_R(A)(x) \leq 0.3$ ). Se puede concluir que los vecinos de los vecinos de  $x$  tienen un grado de pertenencia al subconjunto borroso  $A$ , poco homogéneo y mas alejado del valor  $A(x)$ , según  $TH_{a,R}(A)(x)$  esté más próximo a 1.

**Proposición 5.-** Sea  $R$  una relación reflexiva y simétrica. Sea  $TH_{a,R}(A)$  el gradiente Top-Hat de la apertura. En estas condiciones, se verifica:

- 1)  $TH_{a,R}(A)(x) = 1 \Leftrightarrow A(x) = 1$  y  $\forall z \in R_x \exists h \in R_z \ni A(h) = 0$ .
- 2)  $A(x) \leq 0.5 \Rightarrow TH_{a,R}(A)(x) = A(x)$
- 3)  $A(x) = 0 \Rightarrow TH_{a,R}(A)(x) = 0$  y  $\forall z \in R_x \exists h \in R_z \ni A(h) = 0$
- 4)  $A(x) > 0.5$  y  $A(x) = \gamma_R(A)(x) \Rightarrow TH_{a,R}(A)(x) = 1 - \gamma_R(A)(x) = 1 - A(x)$
- 5)  $A(x) > 0.5$  y  $\gamma_R(A)(x) < A(x) \Rightarrow 1 - A(x) < TH_{a,R}(A)(x) \leq A(x)$

**Proposición 6.-** Sea  $x \in U$  y  $R$  una relación reflexiva y simétrica. Entonces:

- 1) El Top-Hat del cierre se puede escribir,  $TH_{c,R}(A)(x) = \min(\zeta_R(A)(x), 1 - A(x)) \quad \forall x \in U$ .
- 2)  $TH_{c,R}(A)(x) = \alpha$  para  $x \in U$  y  $\alpha \in [0, 1]$  es equivalente a:  $1 - \alpha \geq A(x)$ , existe  $z \in R_x$  tal que  $A(h) \leq \zeta_R(A)(x)$   $\forall h \in R_z$  y  $\forall z \in R_x$  existe  $h \in R_z$  tal que  $A(h) \geq \zeta_R(A)(x)$ .

Observar que  $\zeta_R(A)(x) \geq \alpha$ .

Los resultados de las proposiciones 4 y 6 quedan reflejados en las Figuras 6-7 para nuestro ejemplo.

**Proposición 7.-** Sea  $R$  una relación reflexiva y simétrica. Sea  $TH_{a,R}(A)$  el gradiente Top-Hat de la apertura. Entonces se verifica:

- 1)  $TH_{c,R}(A)(x) = 1 \Leftrightarrow A(x) = 0$  y  $\forall z \in R_x$  existe  $h \in R_z$  tal que  $A(h) = 1$
- 2)  $TH_{c,R}(A)(x) = 0 \Leftrightarrow A(x) = 1$  ó existe  $z \in R_x$  tal que  $A(h) = 0 \quad \forall h \in R_z$
- 3)  $A(x) \geq 0.5 \Rightarrow TH_{c,R}(A)(x) = 1 - A(x)$
- 4)  $A(x) < 0.5$  y  $A(x) = \zeta_R(A)(x) \Rightarrow TH_{c,R}(A)(x) = \zeta_R(A)(x) = A(x)$
- 5)  $A(x) < 0.5$  y  $A(x) < \zeta_R(A)(x) \Rightarrow A(x) < TH_{c,R}(A)(x) \leq 1 - A(x)$ .

Como ejemplo, en la Tabla 5, se dan los datos descriptivos correspondientes al 0.75-corte ( $\alpha = 0.75$ ) de los conjuntos Top-Hat de la apertura y del cierre del conjunto borroso  $A$  asociado a la relación  $R$ . Pueden observarse diferencias apreciables respecto a los datos de  $A$  y entre a los datos de  $TH_{a,R}(A)_{0.75}$  y  $TH_{c,R}(A)_{0.75}$ .

En la definición del Top-Hat podemos utilizar la t-norma  $t(a, b) = \max(0, a+b-1)$  para caracterizar la intersección borrosa, obteniéndose los siguientes resultados.



**Proposición 8.-** Sea  $x \in U$  y  $R$  una relación reflexiva y simétrica. Entonces:

- 1) El Top-Hat de la apertura,  $TH_{a,R}(A)(x) = A(x) - \gamma_R(A)(x) \quad \forall x \in U$ .
- 2)  $TH_{a,R}(A)(x) = \alpha$  para  $x \in U$  y  $\alpha \in [0,1]$  es equivalente a: existe  $z \in R_x$  tal que  $A(h) \geq A(x) - \alpha \quad \forall h \in R_z$  y  $\forall z \in R_x$  existe  $h \in R_z$  tal que  $A(h) \leq A(x) - \alpha$ .

**Proposición 9.-** Sea  $R$  una relación reflexiva y simétrica. Sea  $TH_{a,R}(A)$  el gradiente Top-Hat de la apertura, entonces:

- 1)  $0 \leq TH_{a,R}(A)(x) \leq A(x)$
- 2)  $TH_{a,R}(A)(x) = 0$  para  $x \in U$  es equivalente a:  $A(x) = \gamma_R(A)(x) \Leftrightarrow$  existe  $z \in R_x$  tal que  $A(h) \geq A(x) \quad \forall h \in R_z$  y  $\forall z \in R_x$  existe  $h \in R_z$  tal que  $A(h) \leq A(x)$
- 3)  $TH_{a,R}(A)(x) = 1 \Leftrightarrow A(x) = 1$  y  $\forall z \in R_x$  existe  $h \in R_z$  tal que  $A(h) = 0$ .
- 4)  $TH_{a,R}(A)(x) = A(x) \Leftrightarrow \gamma_R(A)(x) = 0 \Leftrightarrow \forall z \in R_x$  existe  $h \in R_z$  tal que  $A(h) = 0$ .

**Proposición 10.-** Sea  $x \in U$  y  $R$  una relación reflexiva y simétrica. Entonces:

- 1) El Top-Hat del cierre,  $TH_{c,R}(A)(x) = \zeta_R(A)(x) - A(x) \quad \forall x \in U$ .
- 2)  $TH_{c,R}(A)(x) = \alpha$  para  $x \in U$  y  $\alpha \in [0, 1]$  es equivalente a: existe  $z \in R_x$  tal que  $A(h) \leq A(x) + \alpha \quad \forall h \in R_z$  y  $\forall z \in R_x$  existe  $h \in R_z$  tal que  $A(h) \geq A(x) + \alpha$ .

Los resultados de las proposiciones 9-10 quedan reflejados en las Figuras 8 y 9.

**Proposición 11.-** Sea  $R$  una relación reflexiva y simétrica. Sea  $TH_{a,R}(A)$  el gradiente Top-Hat de la apertura. Entonces:

- 1)  $0 \leq TH_{c,R}(A)(x) \leq \zeta(x)$
- 2)  $TH_{c,R}(A)(x) = 0 \Leftrightarrow A(x) = \zeta_R(A)(x) \Leftrightarrow$  existe  $z \in R_x$  tal que  $A(h) \leq A(x) \quad \forall h \in R_z$  y  $\forall z \in R_x$  existe  $h \in R_z$  tal que  $A(h) \geq A(x)$
- 3)  $TH_{c,R}(A)(x) = 1 \Leftrightarrow A(x) = 0$  y  $\forall z \in R_x$  existe  $h \in R_z$  tal que  $A(h) = 1$
- 4)  $TH_{c,R}(A)(x) \geq 0,5 \Leftrightarrow \zeta_R(A)(x) \geq 0,5$  y  $A(x) \geq 0,5$
- 5)  $A(x) = 0 \Leftrightarrow TH_{c,R}(A)(x) = \zeta_R(A)(x) \Leftrightarrow$  existe  $z \in R_x$  tal que  $\forall h \in R_z A(h) \leq A(x) + \zeta_R(A)(x)$  y  $\forall z \in R_x$  existe  $h \in R_z$  tal que  $A(h) \geq A(x) + \zeta_R(A)(x)$ .

**Análisis del filtro Hit-OR-Miss**

Sea  $R$  una relación binaria ( $R \subseteq U \times U$ ) y sea  $R_1$  una relación tal que  $R_1 \subseteq R$ . Consideramos  $R_2 = R - R_1 = \{(x,y) \in U \times U \mid R(x,y) = 1 \wedge xR_1y=0\}$ .

Tabla 4: Estadísticos descriptivos de  $\gamma_R(A)$

	$\gamma_R(A)$
Media GraPerA	0.6
Media de edad	43,6
Media años-estudio	8,6
Media H-T-S	39.6
Media SalAnu	197270
Porcentaje de mujeres	45,6%
Porcentaje de blancos	34.9%

Tabla 5: Estadísticos de  $TH_{a,R}(A)_{0.75}$  y  $TH_{c,R}(A)_{0.75}$

	$TH_{a,R}(A)_{0.75}$	$TH_{c,R}(A)_{0.75}$
Media GraPerA	0.9	0.04
Media de edad	37	37.4
Media años-estudio	10	10.2
Media H-T-S	40	40.6
Media SalAnu	222762	168129
Porcentaje de mujeres	33.7%	30.7%
Porcentaje de blancos	88.4%	91.8%

En estas condiciones se verifica:

- 1)  $R_1 \cap R_2 = \emptyset$
- 2) Si  $R$  y  $R_1$  son simétricas, entonces,  $R_2$  es simétrica.
- 3) Si las relaciones  $R$  y  $R_1$  son reflexivas, entonces  $R_2$  es no reflexiva.
- 4) Si la relación  $R$  es no reflexiva, entonces  $R_1$  y  $R_2$  son no reflexivas.

Dado un subconjunto  $A \subseteq U$  y dos relaciones reflexivas y simétricas  $R \subseteq U \times U$  y  $R_1 \subseteq R$ , se define el operador Hit-or-Miss.

**Definición 4.-** Sea  $A$  un subconjunto borroso de  $U$ ,  $R$  una relación binaria ( $R \subseteq U \times U$ ),  $R_1$  una relación, tal que  $R_1 \subseteq R$  y  $R_2 = R - R_1 = \{(x,y) \in U \times U \mid xRy=1 \wedge xR_1y=0\}$ . Se define el *Hit-Or-Miss* de  $A$ , asociado al par de relaciones  $(R_1, R_2)$ , por la siguiente expresión:

$HM_{(R_1,R_2)}(A) = E_{R_1}(A) \cap E_{R_2}(A^C),$

(10)

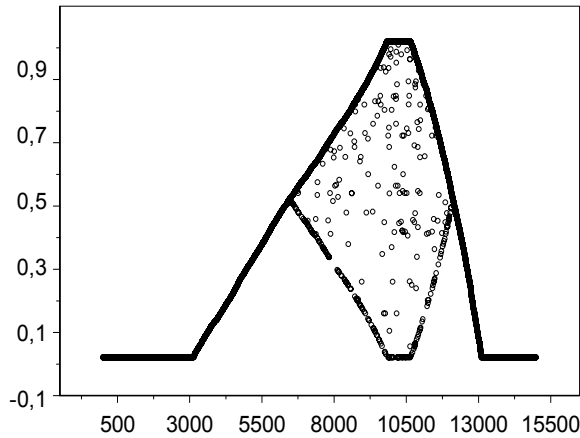


Fig. 6: Grafo de  $TH_{a,R}(A)(x)$

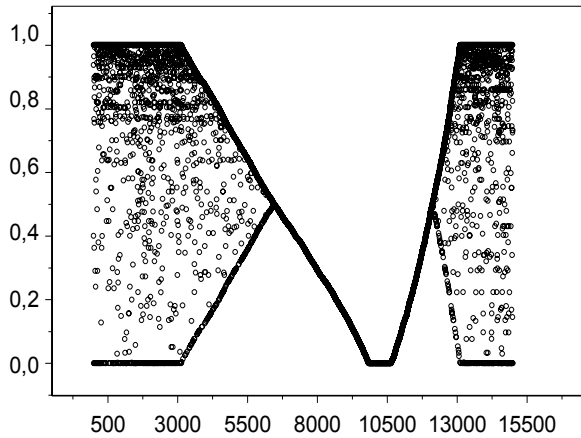


Fig. 7: Grafo de  $TH_{c,R}(A)$

que se puede expresar mediante la función de pertenencia  $HM_{(R_1,R_2)}(A)(x) = \min(E_{R_1}(A)(x), E_{R_2}(A^c)(x))$

*Proposición 12.-* En las condiciones de la definición 4, y considerando que  $(E_{R_1}(A) \cap E_{R_2}(A^c))(x) = \min(E_{R_1}(A)(x), E_{R_2}(A^c)(x))$ , se tiene:

- 1)  $HM_{(R_1,R_2)}(A)(x) = \alpha$  para  $x \in U$  y  $\alpha \in [0, 1]$  es equivalente a:  $[A(z) \geq \alpha$  para todo  $z \in R_{1x}$  y existe  $z \in R_{1x}$  tal que  $A(z) = \alpha$  y  $1 - \alpha \geq A(z)$  para todo  $z \in R_{2x}$ ] o  $[A(z) \geq \alpha$  para todo  $z \in R_{1x}$  y existe  $z \in R_{1x}$  tal que  $1 - \alpha = A(z)$  y  $1 - \alpha \geq A(z)$  para todo  $z \in R_{2x}$ ].
- 2)  $HM_{(R_1,R_2)}(A)(x) = A(x)$  si y sólo si  $A(z) \geq A(x) \ \forall z \in R_{1x}$  y  $1 - A(x) \geq A(z) \ \forall z \in R_{2x}$ .
- 3)  $HM_{(R_1,R_2)}(A)(x) = 0$  si y sólo si existe  $z \in R_{1x}$  tal que  $A(z) = 0$  o existe  $z \in R_{2x}$  tal que  $A(z) = 1$ .
- 4)  $HM_{(R_1,R_2)}(A)(x) = 1$  si y sólo si  $A(z) = 1$  para todo  $z \in R_{1x}$  y  $A(z) = 0$  para todo  $z \in R_{2x}$ .

Para los datos del ejemplo y utilizando las relaciones definidas por:

$R(x,y) = 1 \Leftrightarrow |x.Ca5 - y.Ca5| \leq 1 \text{ AND } x.Ca1 = y.Ca1 \text{ AND } x.Ca9 = y.Ca9$ . En otro caso  $R(x,y) = 0$ .

$R_1(x,y) = 1 \Leftrightarrow |x.Ca5 - y.Ca5| \leq 1 \text{ AND } x.Ca1 = y.Ca1 \text{ AND } x.Ca9 = y.Ca9 \text{ AND } x.Ca13 = y.Ca13$ .

En otro caso  $R_1(x,y) = 0$ .

$R_2(x,y) = 1 \Leftrightarrow |x.Ca5 - y.Ca5| \leq 1 \text{ AND } x.Ca1 = y.Ca1 \text{ AND } x.Ca9 = y.Ca9 \text{ AND } x.Ca13 \neq y.Ca13$ .

En otro caso  $R_2(x,y) = 0$ .

Se tiene que de las 15000 persona registradas en  $U$ , sólo 182 personas han obtenido  $HM_{(R_1,R_2)}(A)(x) > 0.5$ , y de ellas 21 han obtenido  $HM_{(R_1,R_2)}(A)(x) = 1$ .

Por la proposición 12 se puede decir que:

- 1)  $HM_{(R_1,R_2)}(A)(x) = 1$  si y sólo si todas las personas con la misma edad y que pertenecen al mismo grupo étnico que  $x$ , tales que la diferencia de años de estudio con respecto a  $x$  sea menor o igual que 1, tienen 1 ó 0 como grado de pertenencia al subconjunto  $A$ , según coincida o no el número de horas de trabajo semanal con el de  $x$ .
- 2)  $HM_{(R_1,R_2)}(A)(x) = 0$  si y sólo si existe una persona con la misma edad, y grupo étnico que  $x$ , cuya diferencia de años de estudio con respecto a  $x$  es menor o igual que 1, tal que, o trabaja las misma horas semanales que  $x$  y su grado de pertenencia al subconjunto  $A$  es 0, ó trabaja un número de horas semanales distinto que  $x$  y su grado de pertenencia al subconjunto  $A$  es 1.

En la figura 10 se representa la distribución de los valores  $HM_{(R_1,R_2)}(A)(x)$ , con los datos del Ejemplo y tomando las relaciones binarias anteriores. En la Tabla 6, se dan los datos descriptivos correspondientes al subconjunto  $HM_{(R_1,R_2)}(A)$  del subconjunto borroso  $A$ , pueden apreciarse diferencias notables, respecto a los datos de  $A$  y el resto de conjuntos analizados.

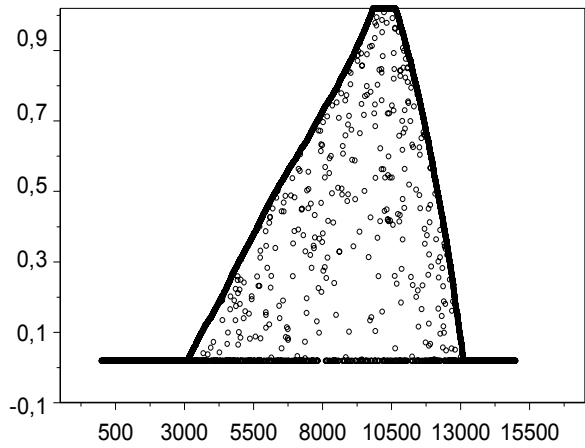


Fig. 8: Grafo de  $TH_{a,R}(A)$ , si se utiliza la  $t$ -norma  $t(a, b) = \max(0, a+b-1)$  para caracterizar la inclusión borrosa.

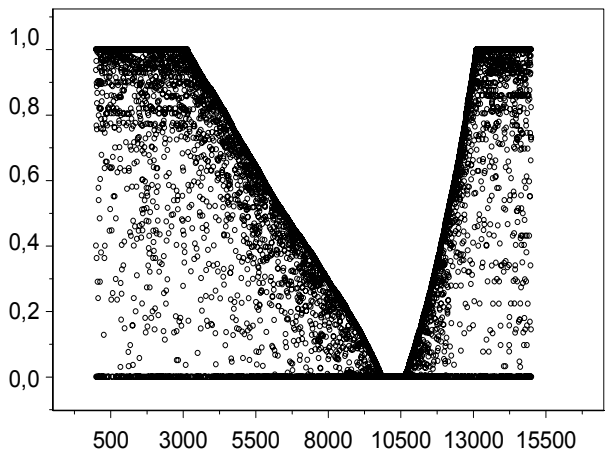


Fig. 9: Grafo de  $TH_{c,R}(A)$  si se utiliza la  $t$ -norma  $t(a, b) = \max(0, a+b-1)$  para caracterizar la inclusión borrosa.

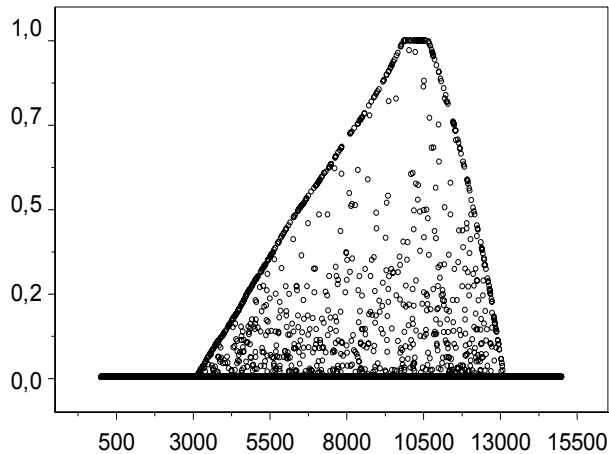


Fig. 10: Grafo de la función  $HM_{(R_1,R_2)}(A)(x)$

Tabla 6: Estadísticos de $HM_{(R_1, R_2)}(A)$	
	$HM_{(R_1,R_2)}(A)$
Media GraPerA	0.55
Media de edad	45,6
Media años-estudio	9,2
Media H-T-S	40,3
Media SalAnu	195042
Porcentaje de mujeres	35,4
Porcentaje de blancos	29,1

CONCLUSIONES

En este trabajo se han aplicado métodos de la Morfología Matemática Borrosa al proceso de filtrado de Datos, considerando como referencial una tabla  $U$ , un subconjunto borroso  $A \in [0,1]^U$  y una relación  $R \in [0,1]^{U \times U}$  como elemento estructurante.

Dado un referencial  $U$ , un subconjunto borroso  $A$  de  $U$  y una relación ordinaria  $R \in \{0,1\}^{U \times U}$ ; la Morfología Matemática permite realizar distintos tipos de filtrado, obteniendo subconjuntos borrosos de  $U$ ,  $E_R(A)$ ,  $\gamma_R(A)$ ,  $TH_{a,R}(A)(x)$ , etc. El grado de pertenencia de un elemento  $x \in U$  a uno de estos subconjuntos se determina en función de la verificación o no, por parte de los elementos de su  $R$ -vecindad o  $R^{op}$ -vecindad, de condiciones preestablecidas por la operación seleccionada para el filtrado.

El interés de estos filtros radica en que para cada uno de los elementos del subconjunto obtenido del filtrado, además de las características de ese elemento, disponemos de información sobre los elementos de sus respectivas  $R$ -vecindades, información que depende del operador utilizado. Los elementos seleccionados en  $U$  por un filtro, pueden considerarse como patrones de grupos de ítems con características preestablecidas por la relación y la operación utilizadas, así como por el subconjunto borroso  $A$  con el que estamos trabajando.

Los operadores utilizados en este trabajo se implementaron en una base de datos ACCESS, se utilizaron las sentencias clásicas de SQL para la selección de registros en la programación de los

algoritmos. Se considera que estos métodos pueden integrarse sin dificultad en lenguajes de bases de datos que incorporan el lenguaje SQL, y por tanto, el comportamiento con grandes volúmenes de información va a depender fundamentalmente de la base de datos y del lenguaje de programación que incorpore.

El objetivo de este trabajo estaba limitado a investigar las posibilidades de utilizar los métodos de la Morfología Matemática Borrosa en la búsqueda de información relevante en una base de datos. Por ello en este trabajo no aparece un estudio comparativo con otros métodos. Se puede decir que los operadores dilatación, erosión, apertura y cierre de un subconjunto borroso  $A$  de  $U$  tienen un coste computacional de orden  $\theta(n^2)$ , siendo  $n$  el número de registros del referencial  $U$ .

Como se ha comprobado con el ejemplo analizado en este trabajo, el análisis de los subconjuntos borrosos de registros seleccionados y la comparación de estadísticos con el subconjunto borroso  $A$  y el referencial  $U$ , puede aportar conocimiento, entendiendo como tal la obtención de nueva información desconocida previamente. Por ello parece interesante seguir investigando sobre el papel que pueden desempeñar los filtros morfológicos en el descubrimiento de conocimientos en las Bases de Datos.

## REFERENCIAS

- Bloch, I. y H. Maître; Fuzzy Mathematical Morphologies: a comparative study, Pattern Recognition, 28, 1341-1387 (1995).
- Codd, E.F.; A Relational Model of Data for Large Shared Data Banks, ACM, 13(6): 377-387 (1970)
- Divyendu, S. y E.R., Dougherty, Characterization of fuzzification Minkowski Algebra, SPIE Vol 1769, Image Algebra and Morphological Image Processing III, 59-69 (1992).
- Erich, P.K., R. Mesiar y E. Pap; Triangular Norms. Kluwer Academic Publishers (2000).
- Frago, N. y R. Fuentes-González; Técnicas de Morfología Matemática en el tratamiento Sistemas Relacionales Difusos. Actas del Congreso Español Sobre Tecnologías y Lógica Fuzzy. Ciudad Real (España), Septiembre (2006).
- Fuentes-González, R., P. Burillo y N. Frago; Aportación al proceso de Descubrimiento de Conocimiento utilizando técnicas de Morfología Matemática. Actas del Congreso Español Sobre Tecnologías y Lógica Fuzzy. León (España), Septiembre (2002).
- Frawley, W.J., G. Piatetsky-Shapiro y C.J. Matheus; Knowledge discovery databases: An Overview. In Piatetsky-Shapiro y Frawley (eds) Knowledge Discovery in Databases, AAAI/MIT, 1-27 (1991).
- Heimans, H. y L. Vincent; Graph Morphology in Image Analysis, in E.R. Dougherty (Ed.) Mathematical Morphology in Image Processing, 171-204, Marcel Dekker, Inc. (1993).
- Matheron, D.; Elements pour une théories des milieux poreus, Masson, Paris (1967).
- Serra, J.; Image Analysis and Mathematical Morphology. Academic Press Vol I (1982) -II (1988)
- Zadeh, L. A.; Fuzzy Sets, Information Control, 8:338-335 (1965).